PATENT
ATTORNEY DOCKET NO.: INGA,004

APPLICATION FOR UNITED STATES LETTERS PATENT

Entitled

PRODUCTION OF ssDNA *IN VIVO*

Applicant: Charles A. Conrad
Over Land Park, Kansas

Assignee: Ingene, Inc.

CERTIFICATE OF MAILING BY EXPRESS MAIL

Express Mail Number: TB895079713US
Date of Deposit: October 9, 1998

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR §1.10 on the date indicated above and is addressed to Box Patent Application, The Honorable Commissioner of Patent and Trademarks, Washington, D.C. 20231.

Mark R. Wisner, Registration No. 30,603 — October 9, 1998
Date of Signature

PRODUCTION OF ssDNA *IN VIVO*

BACKGROUND OF THE INVENTION

This application is a continuation-in-part of my co-pending application Serial No. 08/877,251, entitled STEM-LOOP CLONING VECTOR AND METHOD, filed June 17, 1997. Serial No. 08/877,251 is itself a continuation application of application Serial No. 08/236,504, having the same title, filed April 29, 1994.

The present invention relates to a stable DNA construct, conveniently referred to as a cassette, into which a target nucleic acid strand is incorporated for subsequent replication in a prokaryotic or eukaryotic host cell, and expression within eukaryotic host cells. The construct, or cassette, includes inverted tandem repeats which form the stem of a stem-loop structure which includes one or more genes which function *in vivo* to cause expression of the single stranded DNA target sequence.

So far as is known, there is no method for producing single-stranded deoxyribonucleic acid (ssDNA) species in eukaryotic cells which do not contain intervening and/or flanking vector sequences. The scientific and patent literature does include the disclosure of cDNA-producing vectors (*see* A. Ohshima, *et al.*, 89 Proc. Natl. Acad. Sci. USA 1016-1020 (1992); S. Inouye, *et al.*, 3 Current Opin. Genet. Develop. 713-718 (1993); O. Mirochnitchenko, *et al.*, 269 J. Biol. Chem. 2380-2383 (1994); J.-R. Mao, *et al.*, 270 J. Biol. Chem. 19684-19687 (1995); and U.S. Patent No. 5,436,141), but that system does not appear to have demonstrated the ability to produce ssDNA in eukaryotic cells without intervening vector sequences. The ssDNA produced by the method described in these references includes intervening nucleotide sequences which can interfere with the intended function of the ssDNA product.

There are also a number of viral and transposable elements that have been discovered which contain ssDNA intermediates within their life cycles in eukaryotic and prokaryotic systems (*see* A.M. Weiner, *et al.*, 55 Ann. Rev. Biochem. 631-661 (1986) and H. Varmus, *et al.*, *in* Mobile DNA, M.M. Howe and D.E. Berg (Eds.), American Society for Microbiology: Washington, D.C., pp. 53-108 (1989)). Many of these genetic elements can be adapted to produce single-stranded nucleic acids within eukaryotic systems but would necessarily contain genetic information (nucleotide bases) which could interfere with the desired function of the *in vivo* produced single-stranded

oligonucleotides or are difficult to manipulate beyond their normal biological life cycles to carry desired nucleotide sequences into the cell. The RNA expression vectors which have been described in the literature produce antisense RNA in cells, but the RNA:RNA hybrids which are formed are not as stable as DNA:RNA hybrids, the synthesized RNA has a shorter half-life than DNA, and RNA-RNA hybrids do not stimulate RNase H activity as does the DNA:RNA hybrids (H. Donis-Keller, 7 Nucleic Acids Res. 179-192 (1979)).

Artificially synthesized DNA analog oligomers used for antisense therapies have had a number of difficulties, including problems in cell uptake and distribution (P.A. Cossum, *et al.*, 267 J. Pharmacol. Expl. Ther. 1181-1190 (1993); H. Sands, *et al.*, 47 Mol. Pharmacol. 636-646 (1995)) as well as toxicity problems due to the high blood concentrations required to be effective (S.P. Henry, *et al.*, 116 Toxicology 77-88 (1997)). The DNA analogs which are by far the most used in antisense therapies are phosphorothioates and methylphosphonates. However, phosphorothioate oligonucleotides tend to bind serum and intracellular proteins nonspecifically (S.T. Crooke, *et al.*, 227 J. Pharmacol. Exp. Ther. 923-937 (1996); W.Y. Gao, *et al.*, 41 Mol. Pharmacol. 223-229 (1992)), and at higher concentrations, inhibit RNase H activity (S.T. Crooke, *et al.*, 312 Biochem. J. 599-608 (1995)). Phosphorothioate oligonucleotides have a lower Tm (an average of 0.50C per base-pair) for RNA than does DNA (S.T. Crooke, *et al.*, Antisense Research and Application, CRC Press: Boca Raton (1993)), which requires that phosphorothioate oligonucleotides be typically longer than phoshodiester DNA oligonucleotides for effective binding, which can cause a loss of hybridization specificity (J.-J. Toulmé, *et al.*, *in* C. Lichtenstein and W. Nellen (Eds.), Antisense Technology: A Practical Approach IRL Press: New York, pp. 39-74 (1997)). Further, although their pharmacokinetic properties appear to be more favorable than unmodified oligonucleotides, the half life of phosphorothioate oligonucleotides in many experimental animals is less than one hour (S. Agrawal, *et al.*, 88 Proc. Natl. Acad. Sci. USA 7595-7599 (1991); P. Iverson, 6 Anticancer Drug Design 531-538 (1991)) such that delivery of an efficacious dose of the oligonucleotide may be problematical. Methylphosphonate oligonucleotides do not activate RNase H enzyme activity (L.J. Maher, *et al.*, 245 Science 725-730 (1989); P.S. Miller, *in* J.S. Cohen (Ed.), Oligodeoxynucleotides: Antisense Inhibitors of Gene Expression, CRC Press: Boca

Raton, p, 79 (1989)) and are eliminated rapidly (T.L. Chen, *et al.*, 18 Drug Metab. Dispos. Biol. Fate. Chem. 815 (1990)).

Another factor influencing the delivery of synthetic oligonucleotides into cells is the low permeability of the cell membrane to such compounds. Low permeability may effectively preclude adequate uptake for sustained *in vivo* activity. Consequently, many oligonucleotides must be delivered to the cell by carrier systems such as liposomes or molecular complexing agents. However, the relatively short circulating time of liposomes, lack of targeting specificity, and the need for repeat exposure severely limits the usefulness of synthetic oligonucleotide delivery schemes. These difficulties can be avoided, however, if production of the desired single-stranded oligonucleotide takes place within the cell (*in vivo*).

Nevertheless, so far as is known, no method for producing single-stranded nucleic acid in eukaryotic cells is available which overcomes all of these limitations and disadvantages. It is, therefore, and object of the present invention to provide a method which overcomes these limitations and disadvantages, and in more detail, it is an object of the present invention to provide a DNA construct which directs the synthesis of ssDNA *in vivo*. It is another object of the present invention to provide a method which reduces and/or eliminates the problems due to toxicity, specificity, and Rnase H inhibition or inactivation of prior systems by a continuous process which synthesizes naturally-occurring phosphodiesterase single-stranded DNA within the target cell.

Another object of the present invention is to provide a method, and a DNA construct, for producing ssDNA of any desired nucleotide sequence within eukaryotic cells without undesirable intervening or flanking nucleotide bases.

Another object of the present invention is to provide a method, and a DNA construct, for producing ssDNA *in vivo* in a manner which isolates or presents a sequence of interest to the desired target without interference from intervening or flanking nucleotide bases.

Another object of the present invention is to provide a method, and a DNA construct, for producing ssDNA of any nucleotide sequence *in vivo* which can be used for (but is not limited to) binding to mRNA in an anti-sense fashion to down regulate a gene product or a viral gene product of interest.

Another object of the present invention is to provide a method, and a DNA construct, for producing ssDNA designed in such a way as to favor binding to duplex (native DNA) to form triplex structures which may interfere with normal gene transcription and regulation.

Another object of the present invention is to produce ssDNA within eukaryotic cells for the purpose of disrupting one or more of the many highly regulated cell functions. For instance, the ssDNA tails of telomeric repeats may be altered by the production of ssDNA which has identical or complimentary nucleotide base composition to the sequence of the native DNA in the telomeric repeats or other regulatory sequence.

Another object of the present invention to provide a method, and a DNA construct, for producing ssDNA *in vivo* which is designed so that the nucleotide sequence is recognized by a cellular protein of interest and subsequently binds to and inhibits a specific cellular function, for instance, by binding to proteins which recognize nucleic acid sequences.

Yet another object of the present invention is to provide a method, and a DNA construct, for producing ssDNA into which secondary structures are designed so that the ssDNA oligonucleotides bind to and/or otherwise inhibit or activate various cellular functions which rely on nucleic acid protein interaction such as transcription, translation, and DNA replication.

Another object of the present invention is to provide a method, and a DNA construct, for producing ssDNA *in vivo* for site-directed mutagenesis or gene knockout for therapeutic applications.

Another object of the present invention is to provide a method, and a DNA construct, for producing ssDNA of precisely defined nucleotide composition which favors site-specific insertion into a genome for therapeutic purposes.

Another object of the present invention is to provide a method, and pharmacologically acceptable compositions, for delivery of antisense sequences to the nucleus of target cells in a manner which produces a thereapeutic effect.

This listing of the objects of the present invention is not intended to be a list of all of the objects of the invention. There are a vast number of other cellular functions which are mediated by the cellular genome which, in the interest of brevity and practicality, are not mentioned here and which are amenable to regulation by *in vivo* production of

ssDNA. For instance, exonucleases digest ssDNA much more aggressively than double-stranded (ds) DNA. Consequently, another object of the present invention is to provide a ssDNA construct, and a method of producing that construct *in vivo*, which is not as susceptible to degradation by native exonucleases in the cell as ssDNA. It can be seen from this illustration that this list of some of the objects of the present invention is provided for purposes of exemplification and is not intended to limit the scope of the present invention.

SUMMARY OF THE INVENTION

These objects, and the many others which will be made apparent to those skilled in the art by the following description of the presently preferred embodiments of the invention, are achieved by providing a cassette, or nucleic acid construct, comprised of a nucleic acid sequence which is comprised of a sequence of interest flanked by inverted tandem repeats, a gene encoding an RNA-dependent DNA polymerase and a gene encoding a restriction endonuclease. The cassette also preferably includes a gene encoding an RNase H and either constitutive or inducible eukaryotic promoter(s)/enhancer(s) for the RNA-dependent DNA polymerase and restriction endonuclease genes. The invention also contemplates that the cassette is incorporated into a plasmid and that the plasmid is incorporated into a suitable host cell.

In another aspect, the present invention comprises a method of producing single-stranded DNA *in vivo* comprising the steps of transcription and translation of a cassette comprising am RNA-dependent DNA polymerase gene and a sequence of interest in a eukaryotic cell and converting the mRNA transcript of the sequence of interest to cDNA with the polymerase produced by the RNA-dependent DNA polymerase gene with simultaneous digestion of the mRNA component template with an RNase H expressed enzyme. The sequence of interest also includes an inverted tandem repeat. The cassette may also include a restriction endonuclease gene which, when transcribed and translated, produces a restriction endonuclease which linearizes the transcript of the sequence of interest by cutting the ss-DNA at a restriction endonuclease site formed when the inverted tandem repeat causes the transcript to form a stem-loop intermediate.

In another aspect, the present invention comprises a method for producing a single-stranded oligonucleotide in a target cell. In one embodiment, this method is intended to deliver an anti-sense sequence. In other embodiments, the method is used to

deliver triplex-forming sequences or sequences which are recognized and bound by specific DNA-binding proteins, or other nucleic acids and/or proteins which function in cellular metabolism and/or replication.

The method comprises the encoding of the oligonucleotide into a complementary sequence of interest in a cassette which includes a gene encoding for an RNA-dependent DNA polymerase which preferably includes an RNase H gene and an inducible or constitutive eukaryotic promoter/enhancer appropriate for that polymerase/RNase H gene. The cassette includes a gene encoding a restriction endonuclease (RE) and, in the preferred embodiment, an appropriate promoter/enhancer for that RE gene. The cassette further comprises an inverted tandem repeat and, when assimilated into the target cell, the cassette (including the sequence of interest and the inverted tandem repeats) is transcribed by the cell under the control of the promoter(s)/enhancer(s). The normal function of the target cell causes the resulting mRNA transcript of the polymerase and RE genes to be translated, providing all that is needed for production of ss-DNA from the mRNA transcript of the sequence of interest. Specifically, the RNA-dependent DNA polymerase produced from the cassette converts the mRNA transcript of the sequence of interest and inverted tandem repeats to ss-cDNA, the ss-cDNA forms a stem-loop intermediate as the nucleotide bases comprising the inverted tandem repeats pair up, and the restriction endonuclease produced from the RE gene in the cassette digests the double-stranded portion of the stem-loop intermediate to "free" the single stranded DNA oligonucleotide from the loop portion of the stem-loop intermediate.

Multiple systems can be used to deliver the cassette to the target cell to direct the synthesis of ssDNA within the cell, including plasmid or plasmid-based vector systems or viral based vector systems, and these systems are adapted for that purpose in accordance with standard delivery techniques currently known to the skilled practitioner. These systems include, but are not limited to, viral based systems such as adenovirus, adenoassociated virus, retroviral vectors, and conjugate vectors using double stranded plasma DNA based transfection systems. Once inside the cell, the cassette is transcribed in the normal course of cell metabolism, producing an mRNA transcript of the sequence of interest that is then converted to cDNA by the reverse transcriptase which is likewise produced by the cell from the reverse transcriptase/RNase H gene included in the cassette under the control of the promoter.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic illustration of the production of a plasmid including the cassette of the present invention.

Figure 2 is a schematic illustration of the production of ss-cDNA in a target cell.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

In this description of the preferred embodiments of the present invention, methods and nucleic acid constructs are described for use in producing single-stranded deoxyribonucleic acid (ss-cDNA) oligonucleotides of virtually any predefined or desired nucleotide base composition *in vivo* in eukaryotic cells with or without flanking nucleotide sequences. Methods and constructs are described which use biological rather than the *in vitro* synthesis of ss-cDNA of desired nucleotide base composition. Because biological, i.e., enzymatic reactions, are used in these methods they are applicable to any *in vivo* system.

A vector (as used herein, the term "vector" refers to a plasmid or modified viral construct used to deliver and manipulate DNA segments of interest) was designed to produce ss-cDNA *in vivo* that contains all the necessary enzymatic functions and signaling instructions to allow the host cell to produce ss-cDNA having a desired sequence (a "sequence of interest"). Described herein are the components included in the vector for synthesizing ss-cDNA *in vivo*; they include (1) an RNA-dependent DNA polymerase gene, (2) a restriction endonuclease gene, and (3) a genetic element that supplies the template for the desired cDNA which includes the sequence of interest. Although not required for proper function of the cassette, for reasons set out below, the cassette preferably includes an RNase H gene and appropriate promoter(s)/enhancer(s). Also described herein is a method to construct a vector including these components. Those skilled in the art will recognize from this description, however, that in its most essential form, all that is needed for production of the ss-cDNA sequence of interest in the cell is a nucleic acid construct which can be cut out of the plasmid for subsequent processing and presentation to the cell and that this construct is conveniently referred to as a cassette. The "cassette" is, therefore, that portion of the vector which includes these three elements.

Regarding the RNA-dependent DNA polymerase, or reverse transcriptase (RT) gene which is the first component of the cassette, the reverse transcriptase/RNAse H

gene from Moloney murine leukemia virus was used to advantage in the examples described herein. The reverse transcriptase/RNAse H gene from the human immunodeficiency virus (HIV) was also tested. Many other retroviral reverse transcriptase/RNAse H genes may be used to advantage in the cassette of the present invention, it being preferred that the reverse transcriptase/RNase H gene be a reverse transcriptase/RNase H gene that is regulated by an appropriate upstream eukaryotic promoter/enhancer such as the CMV or RSV promoter for expression in human cells.

Many RNA-dependent DNA polymerase/reverse transcriptase genes are known which are suitable for use in connection with the cassette of the present invention including those from retroviruses, strains of hepatitis B, hepatitis C, bacterial retron elements, and retrons isolated from various yeast species. As found in nature, these RNA-dependent DNA polymerases usually have an associated RNase H component enzyme within the same coding transcript. However, the cassette of the present invention need not include the naturally-occurring RNase H gene for a particular reverse transcriptase. In other words, those skilled in the art will recognize from this disclosure that various combinations of reverse transcriptase and RNase H genes can be spliced together into the cassette of the present invention to fulfill this function, that modifications and/or hybrid versions of these two enzyme systems are available and/or known to those skilled in the art which will function in the intended manner, and/or that the target cell may itself have sufficient endogenous RNase H to fulfill this function.

Those skilled in the art who have the benefit of this disclosure will also recognize that a number of tissue-specific or wide spectrum promoters/enhancers, or combinations of promoters/enhancers other than those listed above may also be used to advantage to control the reverse transcriptase/RNAse H gene, the RE gene, and the sequence of interest. Although a list of all available promoters/enhancers is not needed to exemplify the invention, as noted above, the promoters/enhancers may be constitutive or inducible and may include the CMV or RSV promoters/enhancers listed here and many other viral or mammalian promoters. Representative promoters/enhancers which are appropriate for use in connection with the cassette of the present invention may include, but are not limited to, HSVtk (S.L. McKnight, *et al.*, 217 Science 316 (1982)), human ß-globulin promoter (R. Breathnach, *et al.*, 50 Ann. Rev. of Biochem. 349 (1981)), ß-actin (T. Kawamoto, *et al.*, 8 Mol. Cell Biol. 267 (1988)), rat growth hormone (P.R. Larsen, *et al.*,

83 Proc. Natl. Acad. Sci. U.S.A. 8283 (1986)), MMTV (A.L. Huang, *et al.*, 27 Cell 245 (1981)), adenovirus 5 E2 (M.J. Imperiale, *et al.*, 4 Mol. Cell. Biol. 875 (1984)), SV40 (P. Angel, *et al.*, 49 Cell 729 (1987)), α-2-macroglobulin (D. Kunz, *et al.*, 17 Nucl. Acids Res. 1121 (1989)), MHC class I gene H-2kb (M.A. Blanar, *et al.*, 8 EMBO J. 1139 (1989)), and thyroid stimulating hormone (V.K. Chatterjee, *et al.*, 86 Proc. Natl. Acad. Sci. U.S.A. 9114 (1989)).

The reverse transcriptase/RNase H gene also preferably includes a downstream polyadenylation signal sequence so that the mRNA produced from the reverse transcriptase/RNase H gene includes a 3′ poly(A) tail for mRNA stability. As known to those skilled in the art, multiple poly(A) tails are available and are routinely used for production of expressed eukaryotic genes.

The reverse transcriptase produced in the cell synthesizes a complementary DNA (cDNA) using as the template the genetic element including the sequence of interest described below. The RNase H activity of the reverse transcriptase degrades the mRNA template component of the RNA/cDNA hybrid to produce ss-DNA *in vivo*.

The gene encoding the restriction endonuclease may be any of several genes which encode for restriction endonucleases, and preferably those that are controlled by one or more constitutive or inducible wide spectrum and/or tissue-specific promoters/enhancers such as those listed above. The particular restriction endonucleases tested were MboII and FokI, but those skilled in the art who have the benefit of this disclosure that any restriction endonuclease (type I, II, IIS, or III) site may be included in the inverted tandem repeat. These enzymes "clip" or digest the stem of the stem-loop intermediate described below to linearize the sequence of interest as single-stranded DNA.

Expression of this second enzyme is also regulated by an appropriate constitutive or inducible promoter/enhancer located upstream from the restriction endonuclease gene such as the CMV or RSV promoter for expression in human cells. Those skilled in the art who have the benefit of this disclosure will also recognize that several other promoters/enhancers such as those listed above may be used to advantage to control the RE gene just as various promoters/enhancers are available for the reverse transcriptase/RNase H gene as discussed above. The RE gene also preferably includes a

downstream polyadenylation signal sequence so that the mRNA transcript from the restriction endonuclease gene will have a 3′ poly(A) tail.

The third component included in the cassette encodes a nucleic acid sequence that provides the template for synthesis of ss-cDNA in target cells. It is this element which includes the sequence of interest and the inverted tandem repeats. As is the case for the above genes, this genetic element is preferably regulated by an appropriate wide spectrum or tissue-specific promoter/enhancer, such as the SV-40 promoter, or combination of promoters/enhancers, located upstream of the genetic element. Also as was the case for the above genes, the promoter/enhancer can either be constitutive or inducible promoter. Those skilled in the art who have the benefit of this disclosure will recognize that, as noted above, a number of other eukaryotic promoters may be used to advantage to control expression of the sequence of interest including CMV, RSV (non-cell type specific) or GFAP (tissue specific).

For expression in eukaryotic cells, the sequence of interest also contains a downstream polyadenylation signal sequence so that the mRNA produced by the sequence of interest has a poly(A) tail. Between the 3′ inverted tandem repeat and the polyadenylation signal, the genetic element contains a primer-binding site (PBS) for initiation of priming for cDNA synthesis. The PBS is a sequence which is complementary to a transfer RNA (tRNA) which is resident within the eukaryotic target cell. In the case of the mouse Maloney reverse transcriptase described herein as being utilized in conjunction with the present invention, the PBS takes advantage of the lysine tRNA. The PBS included within the presently preferred cassette described herein was taken from the actual 18 nucleotide sequence region of mouse Maloney virus. *See* 293 Nature 81. In the case of the reverse transcriptase gene from human immunodeficiency virus also described above, the PBS used was taken from the nucleotide sequence of HIV. Y. Li, *et al.*, 66 J. Virology 6587-6600 (1992). In short, any PBS that is matched to the reverse transcriptase which comprises the cassette may be utilized for this purpose. Multiple sequences of interest, each with its corresponding promoter/enhancer, polyadenylation signal, and PBS, flanked between inverted tandem repeats, and PBS, can be included in the cassette of the present invention if desired, for example, for use in delivering anti-sense sequences to the target cell.

The RNA transcript encoded by the genetic element acts as the template used by the transcriptase/RNAse H and restriction endonuclease enzymes described above to synthesize and process the sequence of interest, which as noted above, can be any desired ss-cDNA. The RNA transcript contains a primer binding site (PBS) that is exclusively bound by the primer tRNA, e.g., uncharged lystRNA for the reverse transcriptase from Moloney murine leukemia virus used in the examples described herein, which the reverse transcriptase recognizes and uses to initiate ss-cDNA synthesis. The PBS which is included in the cassette is the PBS which is appropriate for the particular reverse transcriptase which is utilized.

The cassette of the present invention also comprises a pair of inverted tandem repeats, which cause the ss-cDNA to fold back upon itself to form the stem of a stem-loop structure in the manner described in my co-pending application Serial No. 08/877,251, the specification of which is hereby incorporated into this application in its entirety by this specific reference, after the cassette is transcribed in the cell and after the reverse transcriptase/RNase H produced by transcription of those genes included in the cassette produce the ss-cDNA sequence of interest from the mRNA transcript in the cell. The stem comprises one or more restriction endonuclease site(s) which is cut by the restriction endonuclease produced from the restriction endonuclease gene also is coded for by the cassette also as described in co-pending application Serial No. 08/877,251. The ss-cDNA which is produced is transcribed with the encoded 5' and 3' regions flanking the stem (made up of the inverted repeats) and the loop (containing the sequence of interest). This "stem" structure is comprised of double stranded, anti-parallel DNA and is designed to contain one or more restriction endonuclease recognition sites within the double stranded portion, i.e., the inverted repeats. This stem can also be cut (also termed digested or cleaved) by any number of corresponding restriction endonuclease enzymes which recognize a cut site within the designed stem. The loop portion of the ss-cDNA, which does not form any apparent duplex DNA, is immune to the action of restriction endonuclease activity since restriction endonucleases recognize only double stranded DNA as a target substrate.

One or more restriction endonuclease sites is designed into the inverted repeats which form the stem of the stem-loop intermediate such that the formation of the "stem" of the "stem-loop" intermediate (from the single-stranded transcript) forms a new single

or multiple restriction endonuclease site(s) capable of being cleaved (i.e., because the stem of the stem-loop intermediate is double-stranded). Another option is to design the inverted repeats to contain eukaryotic, prokaryotic, or viral protein DNA binding sites, which can act to competitively titer out selected cellular proteins. Combinations of restriction sites or other genetic elements may be included in the inverted tandem repeats depending on the base pair composition chosen for the construction of inverted repeats such that linear or precisely cut stem-loop intermediate forms of ss-DNA are produced. It is generally preferred to use synthetically constructed genetic elements in the inverted tandem repeats since it is unlikely that a naturally occurring inverted repeat would have the properly aligned restriction sites.

When the cassette is incorporated into a vector, it is preferred that a shuttle vector is used so that the cassette can be amplified on prokaryotic systems and then subsequently expressed in eukaryotic systems. It is preferred that the vector and/or the cassette also include a selective marker gene. For example, amplification of the vector in prokaryotic systems which preferably include ampicillin, kanamycin, or tetracycline resistance genes for positive selection in those systems. For expression in eukaryotic systems, multiple selection strategies may be used including, but not limited to, resistance markers for Zeocin, resistance to G418, or phenotypic selection markers such as *B-gal* or green fluorescence protein.

Incorporation of this cassette into an appropriate vector allows a convenient method to remove predetermined vector sequences after the production of ssDNA. The loop portion of the ssDNA that is produced is comprised of the nucleotide sequence of interest. After digestion with the restriction endonuclease enzyme included in the cassette, the loop is released as linearized, single-stranded cDNA.

It will also be evident to those skilled in the art from this description that the intact stem-loop ss-cDNA structure can function similarly in many applications as the linearized ss-cDNA form. Consequently, the cassette is also used to advantage without the restriction endonuclease gene and associated regulatory elements and/or with a sequence of interest which lacks the corresponding restriction endonuclease site.

It will also be evident to those skilled in the art from this description of the preferred embodiments of the present invention that a cassette can be made which encodes a ss-cDNA that has a "trimmed" stem-loop structure. The restriction

endonuclease sites encoded in the inverted repeats flanking the sequence of interest are designed such that the stem portion (after duplex formation) is digested with the corresponding restriction endonuclease so as to cut the dsDNA comprising the stem in a way that removes a portion of the stem and the associated flanking sequences yet leaves sufficient duplex DNA that the transcript retains the above-described stem-loop structure. Such a ss-cDNA structure may be more resistant to intracellular nucleases by retaining the "ends" of a ssDNA in double stranded form.

It will also be evident to those skilled in the art from this description of the preferred embodiments of the invention that the stem (duplex DNA) can be designed to contain a predetermined sequence that is recognized and bound by specific DNA-binding proteins. Among other uses, such a stem structure is used in the cell as a competitor to titer out a selected protein(s). For example, production of the ss-cDNA stem-loop of the present invention in a cell that contains a binding site for a selected positive transcription factor such as E1a which acts to "bind up" the factor, preventing it from binding a particular promoter and thus inhibiting the expression of a particular deleterious gene. Any desired nucleotide sequence can be inserted into the genetic element which encodes the 'loop' portion to ultimately carry out a desired function, e.g., antisense binding, down regulation of a gene, and so on as herein described.

In another aspect which will be recognized by those skilled in the art, the present invention is used to construct complex secondary ssDNA structures in the loop portion of the DNA transcript produced in accordance with the present invention. Such secondary structure can be engineered to serve any of several functions. For instance, the sequence of interest may include (but is not limited to) a sequence which is incorporated into the loop portion of the single-stranded cDNA transcript which forms so-called "clover leaf" or "crucible" like structures such as those found in the long terminal repeats of adeno-associated virus or in retrotransposons. Under correct circumstances, such structure is integrated in site-specific manner into the host genome.

Because the cassette of the present invention is adaptable for incorporation into multiple commercially available delivery vectors for mammalian and human therapeutic purposes, multiple delivery routes are feasible depending upon the vector chosen for a particular target cell. Such systems include intravenous, intramuscular, and subcutaneous injection, as well as direct intra-tumoral and intra-cavitary injections. The

cassette, when inserted into the vector of choice can also be administered through transmucosal, rectal, oral, or inhalation-type methods of delivery.

The cassette of the present invention is advantageously employed to deliver anti-sense, triplex, or any other single-stranded nucleotide sequence of interest, using known digestion and ligations techniques to splice the particular sequence of interest into the cassette between the inverted tandem repeats. Those skilled in the art who have the benefit of this disclosure will also recognize that the above-described signals used for expression within eukaryotic cells may be modified in ways known in the art depending upon the particular sequence of interest. The most likely change is to change the promoter so as to confer advantageous expression characteristics on the cassette in the system in which it is desired to express the sequence of interest. There are so many possible promoters and other signals, and they are so dependent on the particular target cell for which the sequence of interest has been selected, that it is impossible to list all the potential enhancers, inducible and constitutive promoter systems, and/or poly(A) tailing systems which may be preferred for a particular target cell and sequence of interest.

In one particularly preferred embodiment, the present invention takes the form of a kit comprised of a plasmid having the above-described RNA-dependent DNA polymerase and restriction endonuclease genes cloned therein as well as a cloning site for into which the user of the kit ligates a particular sequence of interest. The cloning site into which the sequence of interest is ligated is located between the above-described inverted tandem repeats. The resulting plasmid is then lyophilized or otherwise preserved for packaging and shipping to the user. The kit preferably also includes the restriction endonuclease for the cloning site into which the sequence of interest is to be cloned and a map of the plasmid along with suitable buffers for ligating the sequence of interest into the cloning site.

Except where otherwise indicated, standard techniques as described by Seabrook, *et al.* (1989) (J. Seabrook, *et al.*, Molecular Cloning: A Laboratory Manual (2nd Ed.), Cold Spring Harbor Press (1989), hereinafter referred to as "Maniatis, *et al.* (1989)") were utilized in the examples set out below. Three experimental designs are presented to illustrate the method of producing ss-cDNA and the process of removing unwanted vector or nucleotide sequences. It should be understood that other methods of production

of ssDNA, both by natural processes and by designed artificial methods using different enzyme products or systems, may also be utilized in connection with the method of the present invention and that the example set out herein are set out for purposes of exemplification as required by the Patent Statute and do not limit the intended scope of this disclosure.

The plasmid pcDNA3.1/Zeo+ was purchased from Invitrogen Corp. (Carlsbad, CA) and plasmid pBK-RSV from Statagene (La Jolla, CA). Oligodeoxynucleotides (ODN) were synthesized by Midland Certified Reagent Co. (Midland, TX). Polymerase chain reactions (PCR) were carried out using Taq DNA polymerase purchased from Boehringer Mannheim Corp. (Indianapolis, IN) in a Robo-gradient thermal cycler (Stratagene (La Jolla, CA). Restriction endonucleases and T4 DNA ligase were obtained from Boehringer Mannheim Corp. (Indianapolis, IN). The ODNs used are listed in the attached Sequence Listing.

EXAMPLE 1. *In vitro* Synthesis of ss-DNA

Four synthetic, single-stranded ODNs (129, 121, 111, and 103 bases in length, Sequence ID Nos. 4, 5, 15, and 16, respectively) were designed (see attached Sequence Listing) to contain an inverted tandem repeat structure with a so-called "stuffer" region of random nucleotide sequence composition inserted as a sequence of interest between an inverted tandem repeat. The inverted tandem repeats were designed in such a way (Fig. 1) as to have a pair restriction endonuclease recognition cut sites within the repeats. The designed cut sites for NotI and FokI (type II and type IIS restriction endonuclease recognition cuts, respectively) were the 111 (Seq. ID 15) and 102 (Seq. ID 16) base length oligonucleotides and the designed restriction endonuclease cut sites for NotI and MboII were the oligonucleotides of length 129 (Seq. ID 4) and 121 (Seq. ID 5) bases. In addition, designed into the oligonucleotides was a tRNA primer binding sequence (PBS) for recognition by primer for reverse transcriptase. The PBS was located 3′ downstream from the inverted tandem repeats.

For each of the synthetic oligonucleotides, 1 μl (5 μg/μl in water) aliquots were added to four separate tubes and heated to 70° for 5 minutes and allowed to self anneal over 15 minutes at room temperature. This process allows optimal hybridization for the formation of a stem-loop structure; the loop portion, which does not have stretches of complementary sequence, should not undergo significant self-annealing and remain

predominantly single-stranded. Standard restriction endonuclease digests were then carried out with NotI, FokI, MboII (appropriate for each oligonucleotide for which a restriction endonuclease cut site was present) and EcoR I (as a negative control) with 10 units of enzyme in a total reaction volume of 15 µl and appropriate reaction buffers. The digests were confirmed by electrophoretic gel analysis.

The results of this experiment showed that a synthetic ss-cDNA with inverted repeats formed duplex DNA. The duplex DNA, presumably the stem of a stem-loop structure, formed specific and recognizable restriction endonuclease sites from the sequences in the inverted tandem repeats, which were designed so as to form NotI and FokI restriction endonuclease sites by Watson-Crick base pairing of the bases comprising the inverted tandem repeats. When the ss-cDNA was incubated with NotI and FokI, the DNA was digested. When incubated with EcoRI (negative control), the same DNA was not digested.

EXAMPLE 2. *In vitro* Formation of ss-DNA From Cassette Transcripts

Two test plasmids were constructed to conduct this experiment. Plasmid pcDNA3.1/Zeo(+) (Invitrogen, Inc.) was digested with NheI and ApaI under standard conditions. Two sets (A and B) of 5' phosphorylated oligonucleotides, which were designed to be complementary with each other (see attached sequence listings) were allowed to hybridize. Hybridization was performed by heating the complementary oligonucleotides at 95°C and then allowing them to cool to room temperature for 15 minutes. The resultant duplex oligonucleotide linkers with appropriate cohesive ends were ligated under standard conditions to the previously prepared pcDNA3.1/Zeo vector.

The selection of positive clones on ampicillin plates was performed after transformation into competent XL1-Blue MRF' cells (Stratagene) as described by Maniatis, *et al.* (1989) and the accompanying instruction. After positive clones were picked, plasmid DNA was isolated using a commercially available plasmid isolation kit (Quiagen, Inc., Sant Caita, CA). Confirmation of DNA ligation was carried out by DNA sequencing.

Positive circular test plasmids were linearized by digestion with PmeI, and standard reverse transcription reactions were carried out as follows: To each tube was added 0.1 µg of linearized plasmid DNA, 25 units of T7 polymerase (i.e., a DNA-dependent RNA polymerase), 2.5 mM rNTP's (ribonucleotide triphosphates rUTP, rCTP,

rGTP, rATP) and appropriate buffer. The reaction tube was incubated at 37°C for 90 minutes. After incubation, 10 units of DNAase was added and incubated for 15 minutes at 37°C. The reaction was terminated by incubation at 70°C for 10 minutes.

Standard cDNA synthesis reactions were carried out using the above reactions. To a fresh tube was added 5 µl of the above reverse transcriptase reaction, 0.2 µg primer (see sequence listing) complementary to the PBS region (downstream from inverted tandem repeats). To this mixture was added 25 units of reverse transcriptase from Moloney murine leukemia virus, 2.5 mM dNTPs (deoxyribonucleotide triphosphates dTTP, dCTP, dGTP, dATP), and the appropriate reaction buffer. The reaction mixtures were incubated at 37°C for 1.5 hours. After the incubation period, 100 units of RNase H was added and the tubes were incubated at 37°C for 15 minutes. The reaction tubes were incubated at 70°C and allowed to anneal by cooling to room temperature over 15 minutes. The mixture was divided equally into four tubes, in which either 10 units of restriction enzyme NotI, FokI, MboII, or EcoRI was added along with the appropriate reaction buffer. The tubes were incubated at 37°C for 1 hour. Replicate reactions were treated with S1 nuclease (specific for ssDNA). The DNA from the above reactions was resolved by gel electrophoresis.

This experiment demonstrated the *in vitro* production of ssDNA by the sequential enzymatic activities of T7 RNA polymerase, reverse transcriptase, and RNase H. The linker contained inverted tandem repeats with the restriction sites NotI, FokI, and MboII; a "stuffer" region (which comprises the loop of the stem-loop intermediate), and a tRNA PBS. The plasmid contained the T7 promoter immediately upstream of the linker region. Agarose gel electrophoresis analysis showed the stepwise production of (1) an mRNA intermediate of predicted length, (2) a cDNA of predicted length, and (3) the subsequent digestion of the ssDNA by the corresponding restriction endonucleases NotI, FokI, and MboII. A negative control in which EcoRI was used did not digest the ss-cDNA.

EXAMPLE 3. *In vivo* Synthesis of ss-cDNA in Eukaryotic Cells

The following *in vivo* experiments were designed to test the plasmids made in Example 2 that contained different "stuffer" regions which were inserted between the inverted tandem repeats. The tissue culture cells used in these experiments encoded and endogenously expressed eukaryotic DNA-dependent RNA polymerase II, which utilizes the RSV promoter located upstream of the cloned genetic element that is the

linker/stuffer segment described in Example 1, above. It will be recognized by those skilled in the art who have the benefit of this disclosure that any eukaryotic promoter can be used for this purpose. It is further recognized by such persons that a vector-encoded DNA-dependent RNA polymerase will also function in this capacity.

Plasmid constructs. The ODNs were allowed to hybridize in 1 μl (5 μg/μl in water) in four separate tubes which were incubated at 70°C for 5 min and allowed to hybridize for 15 min at room temperature. Standard restriction endonuclease digests were carried out (EcoR I used as a negative control) with 10 units of enzyme in a total reaction volume of 15 μl and appropriate reaction buffers. DNA fragments were resolved in and isolated from agarose gels. The selection of positive clones on ampicillin plates was performed after transformation into competent XL1-Blue MRF cells (Stratagene) as described by Maniatis, *et al.* (1989). After positive clones were selected, plasmid DNA was isolated using the above-described Quiagen plasmid isolation kit.

The construction of three sets of expression plasmids is described (Fig.1). The first set of plasmids (Fig. 1A) was derived from pcDNA3.1/Zeo+ (Invitrogen Corp.) and contains the genetic element which encodes the different ss-cDNA sequences used here. pcDNA3.1/Zeo+ was digested with restriction endonucleases NheI and ApaI, which are located in the multiple cloning site (MCS), and where the RT-Not-Mbo-Linker (Seq. ID 4 and Seq. ID 5) was inserted to produce plasmid pc3.1/Zeo/N-M. The RT-Not-Mbo-Linker (Fig. 1B) contains the Mo-MuLV-RT primer binding site (PBS sequence), and within the inverted repeats, the recognition sites for NotI (internal) and MboII (external). Two linker sequences were inserted into the NotI site of the RT-Not-Mbo-Linker. pc3.1/Zeo/N-M was digested with Not I in which was inserted Not-linker-test 1 (Seq. ID 11 and Seq. ID 12) and Not-Mbo-linker-telo (Seq. ID 13 and Seq. ID 14), in which both linkers have compatible NotI ends. The latter linker sequence (Not-Mbo-linker-telo) contains nine repeats of the vertebrate telomere sequence 5'-AGGGTT-3' (E.H. Blackburn, 350 Nature 569-573 (1991)).

The second set of plasmids (Fig. 1C), which contains the Mo-MuLV-RT (T.M. Shinnick, *et al.*, 293 Nature 543-548 (1981)) and restriction endonuclease genes was derived from pBK-RSV (Stratagene). Moloney murine leukemia virus was obtained from the American Type Culture Collection (#CRL-1858). The virus RNA was isolated and prepared for reverse transcriptase-PCR (RT-PCR). A 2.3kb fragment containing the

coding sequence of Mo-MuLV-RT was PCR-amplified using primers as set out in Seq. ID 1 (primer position at nucleotide #2545) and Seq. ID 2 (primer position at nucleotide #4908) to produce a DNA fragment with a 5′-SacI and a 3′-HindIII compatible end. The pBK-RSV was digested with SacI and NheI, the corresponding sites located in the MCS. The NheI end was converted to a HindIII-compatible end by ligating the linker composed of Seq. ID 6 and Seq. ID 7. Insertion of the 2.3 kb M-MuLV-RT gene fragment into pBK-RSV produced pBK-RSV-RT.

The bacterium *Moraxella bovis*, which encodes the restriction endonuclease MboII (H. Bocklage, *et al.*, 19 Nucleic Acids Res. 1007-1013 (1991)), was obtained from the American Type Culture Collection (ATCC#10900). Genomic DNA was isolated from *M. bovis* and used as the template DNA in the PCR. A 1.2kb fragment containing the MboII gene was amplified by PCR using as primers Seq. ID 3 (primer position at nucleotide #887) and Seq. ID 8 (primer position at nucleotide #2206) to produce a fragment with a compatible 5′ HindIII and 3′ XmaI end. The 1.2kb DNA fragment was inserted into pBK-RSV-RT, which had also been digested with HindIII and XmaI, to produce pBK-RSV-RT/Mbo.

A linker sequence encoding alternate histidine and proline amino acids (Seq. ID 9 and Seq. ID 10) was inserted between the RT and MboII genes. pBK-RSV-RT/Mbo was digested with HindIII, and the his-pro linker, with compatible HindIII ends, was inserted at the HindIII site to produce plasmid pBK-RSV-RT/MboL.

The third set of plasmids (Fig. 1D) consists of a fusion of the pc3.1DNA/Zeo-derived plasmids and the pBK-RSV-derived plasmids such that fused plasmids encode the ss-cDNA-encoding genetic element, the Mo-MuLV-RT gene, and the restriction endonuclease (MboII) gene. pBK-RSV-RT/MboL was digested with NsiI to release a 5.3kb fragment containing the Mo-MuLV-RT and MboII genes with the intervening his-pro linker and associated regulatory elements. The 5.3kb DNA fragment was ligated to a linker (Seq. ID 17 and Seq. ID 18) containing an internal EcoRI site. The 5.3kb fragment was digested with EcoRI. The pc3.1/Zeo/N-M and the derivative plasmids containing the test sequence (Seq. ID 11 and Seq. ID 12) and the repeated teleomere sequence (Seq. ID 13 and Seq. ID 14) were digested with BglII, which recognizes a unique site on pc3.1DNA/Zeo in the cytomegalovirus enhancer/promoter (P CMV). The BglII ends were ligated to Seq. ID 17 and Seq. ID 18, which contain the internal EcoRI site. After

digestion with EcoRI, the 5.3kb fragment was ligated to pc3.1/Zeo/N-M and derivatives to generate pcDNAExpress, pcDNAExpress-test, and pcDNAExpress-telo.

Tissue culture studies. Stable and transient transfections were carried out by using lipofectant (Boehringer Mannhiem Corp.) using the manufacturer's accompanying instructions. All plasmid constructs were transfected into Cos-7, U251 and HeLa cell lines. Assays for ssDNA were performed by PCR and by dot-blot analyses 24-48 hours after transfection. Reverse transcriptase activity was assayed using the RT-PCR assay developed by Silver, et al. (Silver, J., *et al.* 21 Nucleic Acids Res. 3593-4 (1993)). The ss-cDNA was isolated from cells transfected 48-72-hr earlier. The ss-cDNA, which co-localizes with RNA, was carried out using trizol reagent (Gibco Life Technologies, Gaithersburg, MD). Assays for specific ss-cDNA species were carried out by both PCR based assays for internal fragment and by denatured single stranded gel electrophoresis with subsequent nylon blotting and probing with an internal biotin-labeled probe.

This experiment showed that human tissue culture cells (HeLa and Cos-7 cell lines), transfected with plasmids designed to synthesize a processed ss-cDNA, produced ss-cDNA of the predicted size. The cells were actually co-transfected with two plasmids, one plasmid carrying the RNA template encoding the stem-loop intermediate (Fig. 1A) and the other carrying the genes for reverse transcriptase and restriction endonuclease (Fig. 1C). Those skilled in the art, however, will recognize that a single plasmid including the RNA template for the stem-loop intermediate and the genes for reverse transcriptase and restriction endonuclease can also be used for this purpose. The studies showed that the plasmid pc3.1/Zeo/N-M (Fig. 1A) synthesized an RNA encoded by the genetic element in Fig. 1B. The genetic element was regulated by the CMV promoter. The synthesized RNA supplied the template for the production of processed ss-cDNA.

The genes encoded in plasmid pBK-RSV-RT/MboL (Fig. 1C) included Moloney murine leukemia virus reverse transcriptase (which also has RNase H activity) and MboII restriction endonuclease. The synthesis of ss-cDNA in eukaryotic cell lines (HeLa) was confirmed using a "test" sequence contained within the linker portion of the plasmid construct pc3.1/Zeo/N-M. The test sequence provides a unique sequence for detection by hybridization using a complementary DNA probe. Using primers complementary to the 5′ and 3′ ends of the synthesized ss-cDNA and RNA isolated from the transfected cell lines (which is where one would expect to find the ss-cDNA fraction), PCR and

agarose gel electrophoresis analyses showed that the ss-cDNA of predicted size was sensitive to S1 nuclease digestion. Appropriate negative and positive controls were included in the PCR analyses. PCR using the same 5′ and 3′ primers and the same isolated RNA amplified DNA of predicted size.

In addition, production of ss-cDNA was demonstrated by dot-blot analysis, in which a probe unique to the test sequence hybridized to the predicted isolates. It was deduced that the ss-cDNA formed a "stem-loop" intermediate *in vivo* because DNA of predicted size was produced if subsequent digestion by MboII had occurred. This latter result shows that the ss-cDNA formed a stem-loop intermediate and that the cloned MboII gene expressed a protein with enzymatic activity.

* * * * *

The experiments described above demonstrate a method of production of ssDNA *in vitro* and *in vivo* by multiple stepwise reactions using eukaryotic reverse transcriptase reactions and various cDNA priming reactions. This reaction was followed by formation of a "stem-loop" intermediate which can be used to eliminate any unwanted sequences either upstream 5′ or downstream 3′ from a designed (and formed) "stem" after being subsequently cleaved by a restriction endonuclease.

Any nucleotide sequence of interest could be produced by this method in a eukaryotic cell. This sequence of interest is cloned (or synthesized) between the designed inverted tandem repeats and represents the sequence in the "loop" after ssDNA production and subsequent stem-loop formation. The sequence of interest to be produced can be of any base (i.e., A,T,G,C) composition as long as the sequence does not interfere with the formation of the stem of the stable stem-loop intermediate, which acts as the substrate for a particular restriction endonuclease. Again, any restriction endonuclease may also be used to digest (or cleave) the stem portion of the stem-loop intermediate as long as the recognition site for that particular restriction endonuclease has been designed into the inverted tandem repeats.

Although described with reference to the figures and specific examples set out herein, those skilled in the art will recognize that certain changes can be made to the specific elements set out herein without changing the manner in which those elements function to achieve their intended respective results. For instance, the cassette described herein is described as being made up of three primary components, a genetic element

which comprises a sequence of interest and a tandem inverted repeat, a restriction endonuclease gene, and a reverse transcriptase gene, each of these genes being provided with appropriate promoters as described herein. Those skilled in the art will recognize that, for instance, the mouse Moloney leukemia virus reverse transcriptase gene described for use as the reverse transcriptase gene of the cassette can be replaced with other reverse transcriptase genes (the reverse transcriptase gene from Human immunodeficiency virus was one such gene which was noted above) and that promoters other than the CMV promoter may be used to advantage. Further, several restriction endonuclease genes are listed above, but those skilled in the art will recognize from this description that the list set out above is not exhaustive and that many other restriction endonuclease genes will function to advantage in connection with the present invention. Similarly, the RSV promoter described as being used in connection with the restriction endonuclease genes set out herein is not the only promoter which may be used to advantage. All such changes and modifications which do not depart from the spirit of the present invention are intended to fall within the scope of the following non-limiting claims.